The Report Committee for Daniel McCravey Justiss
Certifies that this is the approved version of the following report:

A Low Cost Eye-Tracking System for Human-Computer Interface

APPROVED BY
SUPERVISING COMMITTEE:

Supervisor:

Wei Li

Michael Bryant

A Low Cost Eye-Tracking System for Human-Computer Interface

by

Daniel McCravey Justiss

Report

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Science in Engineering

The University of Texas at Austin

August 2017

Abstract

A Low Cost Eye-Tracking System for Human-Computer Interface

Daniel McCravey Justiss, MSE

The University of Texas at Austin, 2017

Supervisor: Wei Li

The quality of life for heavily disable people could be greatly improved if they had a way to interface with a computer. There are already methods in existence attempting to solve this problem, but they all have certain drawbacks. Existing methods are either restricted to use of only more lightly disabled individuals, too expensive, too inaccurate, too invasive, or too uncomfortable to be viable for everyone. Using a system of a laptop with a built-in camera or a computer with a camera otherwise mounted on the screen, it may be possible to create a human computer interface without any of these drawbacks. With image tracking software it can be attempted to determine where on the screen the eyes of the user are looking. Using this data as the input for the computer interface software, the user could control the position of the computer cursor using only their eyes. This method has been attempted several times by several different research groups and no one has managed to develop a system that is accurate and fast enough. However, with one significant change that is proposed in this study it may be possible to have a successful eye tracking system. That new idea is for the user to wear circular eye-glass frames without the lenses that the computer can track along with the eyes. This

would provide additional information to make it possible for an eye tracker of high enough quality for use in a human computer interface. The report describes the design and validation of such a low-cost eye-tracking system for heavily disabled people.

Table of Contents

List of Figures

A Master's Project Report

A Low Cost Eye-tracking System for Human-Computer Interface

By Daniel Justiss
Department of Mechanical Engineering
The University of Texas
Austin, TX 78712

Date:

Chapter 1. Introduction

1.1 Background

There are some people who can only move their eyes. Every other part of their body is paralyzed. Eye movement alone does allow for some very limited communication with other people. For example perhaps looking left and right repeatedly could represent 'no', and looking up and down repeatedly could represent 'yes'. With the development of home computers a potential better way has appeared. It might be possible for a computer to be developed that can be interfaced with using only eye movement. Researchers have been attempting to make this happen for decades. Several different methods have been developed; however, all of them have unresolved problems preventing a viable hands free computer from appearing on the market.

1.1.1 Voice Computer Interface

The two general parts to a voice-interface system is the voice recognition and the application. There are three types of applications: Command and Control, Dictation, and Authentication [1]. Command and Control is about controlling a device with voice commands, such as a computer or a robot. Dictation converts speech into text. Authentication is using a voice to confirm a person's identity, mainly for security.

There are two types of speech recognition systems: speaker dependent and speaker independent. A speaker dependent system needs to be trained by each individual user in order to be accurate, usually by the user reading sentences out loud. A speaker independent system does not need to be trained to be accurate, but cannot be as accurate as a trained speaker dependent system.

There are three components to a speech recognition system: a speech corpus, a frontend processing system, and a speech decoding unit. The speech corpus is a database of words. The frontend processing system extracts features from the input signal which can work as illustrated in Figure 1. The speech decoding unit compares the output of the frontend processing unit to the speech corpus to decide which words were most likely said.

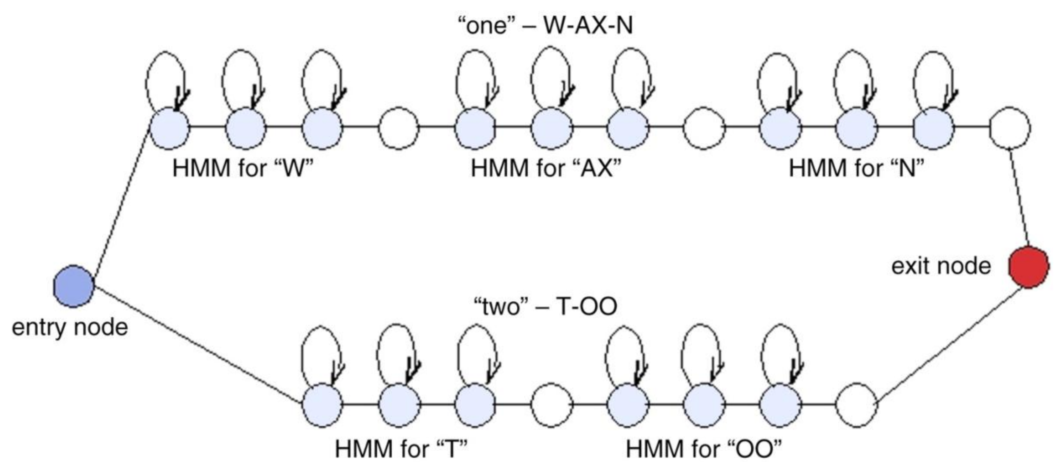

Figure 1. State Diagram of Voice Computer Interface [1]

Figure 1 shows a state diagram of how a computer might recognize particular words. In the case of "two" the letters written are not quite the same as the letters said. Instead two is said something like "too" which has the syllables T and OO. So if a computer recognizes the sounds T and OO being spoken in that order, it will know that the word "two" was spoken as well.

Three commercial speech recognition systems already in existence are Microsoft Speech API version 5 (SAPI5), IBM Via Voice, and Dragon Naturally Speaking.

1.1.2 Neck/face Computer Interface

One method that has been very successful is to use a camera to track a facial feature that the user has muscle control over, such as the nose, eyes, and jaw [2]. When the user rotates their head or moves their jaw, their facial features can be tracked and used to control a computer mouse. The actual facial features can be tracked, or a marker can be placed on the user's face and the marker can be tracked. Such markers might be an LED, an infrared emitter, a mirror reflecting infrared light, or an object brighter than the surroundings. The user could turn their head left, and the cursor would move left. Turn their head up or right and the cursor would go up or right. There are existing systems already that are quite polished, such as Camera Mouse [2].

1.1.3 Brain Computer Interface (BCI)

A BCI interprets signals from the brain into commands that can be received by a personal computer [3,4]. BCI is very attractive because even people with the most serious disabilities can use a BCI. It has three general parts: the sensor(s) receiving data from the brain, the feature extractor, and the feature translator.

The sensor is an electrode recording signals from the neurons of the brain. What these signals mean is not really known. Features from these signals have been observed and there have been attempts to use these signals in BCIs, but the place of these brain wave features in the structure of the brain is simply unknown. It is up to the user to figure out how to excite those brain waves in order to control the BCI. There are three general types of sensors: implants surgically placed inside of the brain, subdural implants which is to say the sensors are placed under the skin against the skull, and electroencephalogram(EEG) which has the sensors placed on the skin. The more invasive the sensor, the better the data quality. The EEG is the cheapest, safest, and most comfortable, therefore naturally the EEG is the best developed.

The feature extractor must extract the desired feature from the raw data. Usually this is done by damping noise and extracting signals from a particular frequency range. The feature translator will then convert this signal into a computer input. It is not scientific from here, whatever commands need an input will be tied to some brain wave feature, and the user will need to learn how to control the computer from there.

1.1.4 Electrooculography

There is a natural magnetic dipole in the eye, with a positive potential in the cornea and a negative potential in the iris [13]. The Electrooculography method is to measure the location of the dipole in order to find the orientation of the eye. The way it works is really quite simple. When two electrodes are placed on either side of the eye, the cornea of the eye disrupts the magnetic field, causing the electrodes to get either a positive or negative reading. The position of the eye determines this reading. This is inexpensive and the measurements are accurate, however it is somewhat obtrusive. Apparently young children detest having the electrodes attached to their face. A commercial example is Eagle Eye [6]. An example is show in Figure 2.

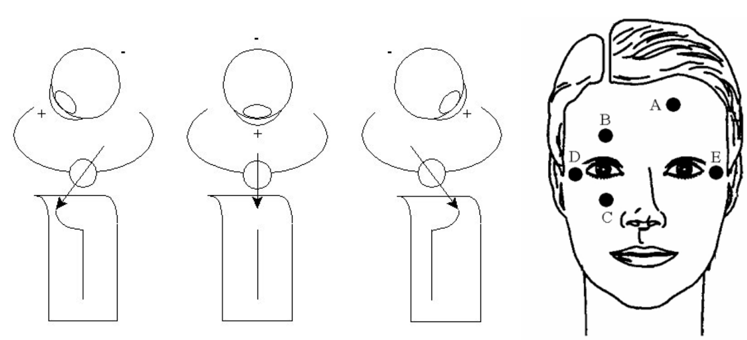

Figure 2. Elecrooculography Diagram [5]

1.1.5 Search Coil

A search coil is a contact lens with some kind of marker placed on it to record eye movements. The two most successful markers are magnetic coils and mirrors [7]. The magnetic coil has a measureable magnetic field when induced with a voltage. The light reflected from the mirror can be used to measure the angle of the eye. Hard contact lenses are typically required to do this, but this causes problems. These lenses are large and obtrusive. The air under the lens needs to be evacuated for the lens to adhere to the eye, which can bruise the eye. Antibiotics also need to be used to prevent infections. The result is fairly accurate however. There is at least one soft lens magnetic coil; however, it cannot sustain prolonged use without drying out the eye and decreasing accuracy [8]. It is also expensive. Overall, search coils are accurate, but uncomfortable and obtrusive.

1.1.6 Camera Tracking Eye Gaze

A digital camera and a computer can be used to translate eye gaze into computer interface inputs [9]. There are many ways to do this, but they all use image detection software to track eye movement. Eye movement alone can be used for very limited commands such as scrolling up or down a page. A potential way to get more complex input to the computer is to use images of the eyes to continuously determine where exactly on the computer screen the user is looking and set the computer cursor to that point. Although not necessarily for the purpose of aiding the disabled, there has been a great deal of research into using cameras to track the point of eye gaze. The most successful method developed so far shines an infrared light onto the user's eyes and uses a special camera that detects infrared light to get images of the eyes with a reflection of the infrared light that was shined onto the eye. Such systems are already on the market with the camera and infrared light emitter built into a headset. One such system is called Tobii Glasses [10]. These systems however are used for market research and not for computer interfacing. The reason for this is unclear; however, it may have something to do with the drawbacks of this sort of systems, which are inaccuracy in environments with natural lighting due to infrared light being in sunlight and high costs of several thousands of dollars. There are efforts to find a way that uses standard cameras that detect red, green, and blue. Such a system would be more accurate in ambient daylight. For this to happen there has to be a frame of reference like the reflected light from the infrared light emitter. The purpose of this frame of reference is to be able to compare the eye images to the location of the computer screen. The most powerful way to do this is to place a camera almost directly in front of each eye, out of center vision just enough that the user can still see the computer screen. Such a system has been developed in the lab of Dr. Aldo Faisal [11]. This system was developed for the express purpose of acting as a computer interface for the disabled. It is quite bulky however and appears to require the user to be laying down and motionless. This does not seem to be a major drawback for a bedridden user however. If the camera were mounted on the computer near the screen this would allow the user to sit rather than laying down. Depending on choices made by the developer it may even allow for the user to be able to move their body a little without destroying the accuracy of the system. One attempt at such a system [12] used a neural network to translate images of the user in front of the computer into computer interface data. This system is fairly accurate though is appears to require the user to limit their movement. A system with a computer mounted camera would be the most convenient and comfortable for the user, however this represents a major challenge for the developers.

1.1.7 Search Mirrors

Search mirrors are notable in that they were the first attempt to track eye gaze, far before even the first digital computers [16]. They used lenses and a mirror embedded into a contact lens on the eye to direct artificial light onto moving rolls of film. There has never been a published attempt to modernize search mirrors using modern equipment. It seems that this method has been abandoned decades ago in favor of other methods.

1.2 Proposed research

The solution proposed in this study is to use a laptop with a built-in camera and a pair of glasses worn by the user that have circular frames and no lenses. Image processing software can find the center of the iris and the center of the circular glasses. These two points when compared to each other can be used to determine where the user is looking on the computer screen. The computer's cursor would move to be where the user is looking. Some cue, perhaps winking, could take the place of clicking the mouse. A virtual keyboard on the screen that could be clicked on would take the place of a physical keyboard. With this, a person who can only use their eyes will be able to communicate more easily, use email, browse the internet, and read every book available online. This would improve such a person's quality of life immensely. This report describes the design and validation of such a system.

1.3 Organization of this report

In the second section there is an in depth review of what has been done on the topic of computer interfacing for disable people. The third section covers the proposed solution for this topic. The fourth section contains the conclusion, reporting the successes and failings of existing solutions and of the proposed solution. There is also discussion on what needs to be done for the proposed solution to be complete.

Chapter 2 Literature Review

2.1 introduction

Of the research topics covered in the last chapter, five are related to the method proposed in this study. They are electrooculography, search coils, search mirrors, infrared, facial recognition, and camera headsets. A brief review of each is provided below.

2.2 electrooculography

The eye has a weak bioelectrical field [13], meaning it can be modeled as a magnet with the positive end centered with the cornea, and the negative end in the back of the eye facing into the skull. The attempt to measure the eye's bioelectric potential for the purpose of calculating eye gaze is called electrooculography(EOG). The electrooculogram has sensors placed around the eyes such as shown in Figure 3. In ideal circumstances as shown in figure 4 the signal corresponds linearly to the position of the eye. The amplitude range of this signal varies with the design of the electrooculogram. There are two signals, one corresponding to the eye in the vertical position, and one corresponding to the signal in the horizontal position. Figure 5 shows the ideal signals corresponding to a simple test shown in Figure 6. The solid line represents the data of vertical eye movement and the dashed line represents the data of the horizontal eye movement. The idea behind the test is for someone equipped with an electrooculogram to follow the markings on the image of Figure 6, in sequence from 1 to 6.

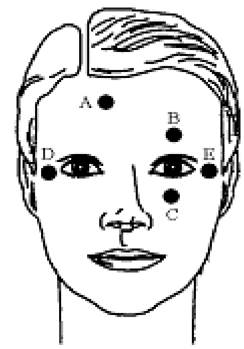

Figure 3. Electrooculography Electrode Array [5]

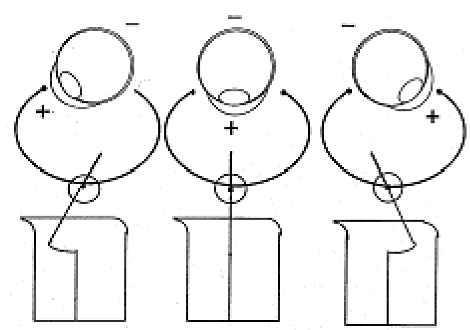

Figure 4. Electrooculography Output Signals Vs. Eye Position [5]

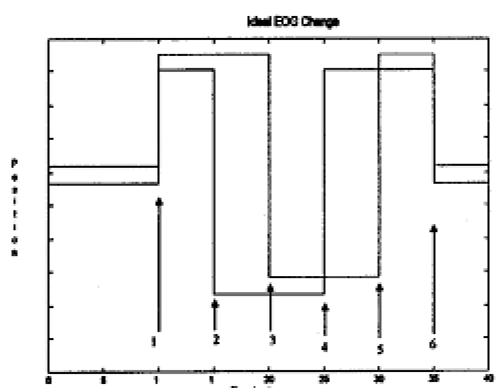

Figure 5. Ideal Electrooculography Signals [13]

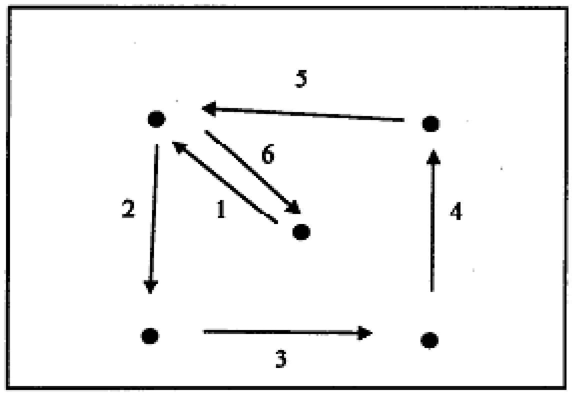

Figure 6. Eye Gaze In Correspondence to Ideal Output Signals [13]

All of this though shows only a perfect, uncorrupted signal. A real signal is corrupted by noise and suffers from a baseline drift. There are also features cause by three biological functions called saccades, fixations, and blinks. Figure 7 showed a real signal compared to its theoretical equivalent. It is easy to see that the real signal is much harder to decipher than the theoretical signal. The standard method of reducing noise is to pass the signal through a low-pass filter. Typical noise is formed by impulses. Impulses are very sharp spikes in the signal which can be described as extremely high frequency waves, and high frequency waves do not pass through low-pass filters. Baseline drift canceling has not been perfected yet [13], but there are well developed methods that work reasonably well.

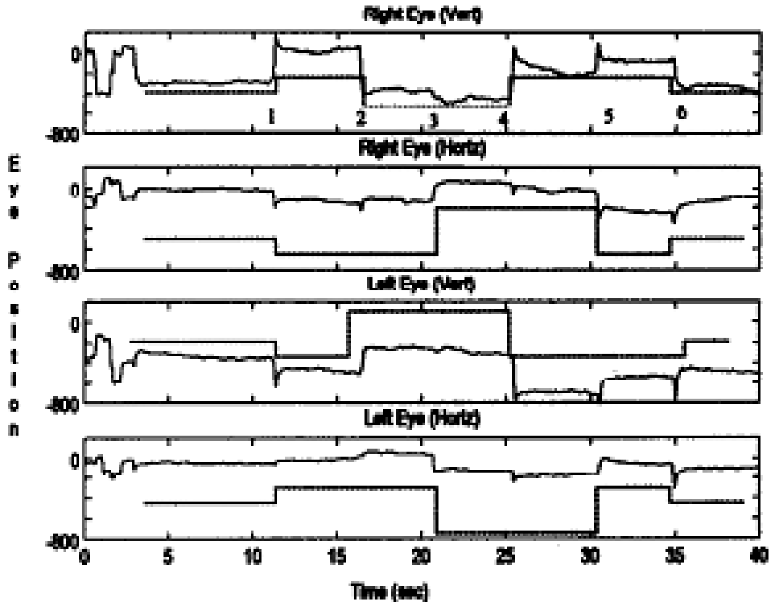

Figure 7. Ideal Signals Vs. Real Signals [13]

Figure 6 shows the features cause by saccades, fixations, and blinks with S representing saccades, F representing a fixation, and B representing blinks. Only a small part of the eye can record in high resolution, so the eye shifts naturally to improve vision quality. These eye shifts are called saccades. Fixations are the times between saccades when the eye is relatively still. Blinks are exactly what they sound like, and they only really distort the vertical signal, and only for a brief moment.

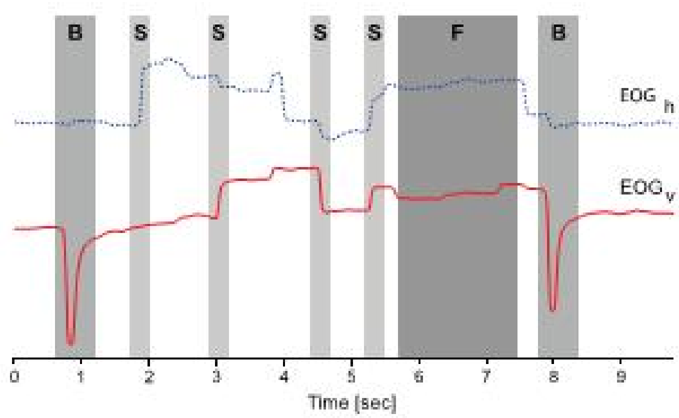

Figure 8. Signal Features [17]

EOG appears to have the potential to be a nearly perfect gaze detector, however after decades of research there is still not a reliable baseline drift filter. With enough effort an EOG might yet can be used to make an accurate computer interface, however it will still have one major drawback. The electrodes stuck to the user's face is very agitating for many people, particularly young children. Eagle Eyes is a computer interface brand using this technology.

2.3Search Coils

A search coil is a coil of wire embedded into a contact lens capable of taking measurements unique to its position in relation to a surrounding magnetic field. The search coil was first developed by David A. Robinson [7] as a tool for psychology research. The system he developed was capable of finding the position of the eye in all 3 degrees of rotation: vertical(θ), horizontal(ϕ), and torsional(ψ). There are other search coil designs, but probably all of them are improvements on Dr. Robinson's original design as shown in figure 9. As such, this is the best place to start.

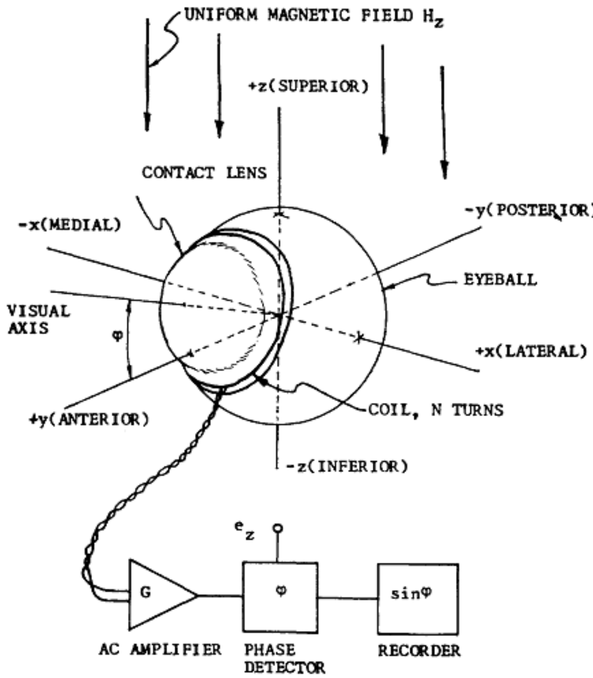

Figure 9. Search Coil Schematic [7]

The only input signal to this system is the alternating magnetic field with gauss peak(H_Z). The only measurable output signal is the voltage of the coil(e). From these variables the vertical angle(ϕ) of the eye can be calculated.

2.4 Search Mirror

There have been scientific efforts to record eye movements as early as the 18th century, but the first method to use photography used a mirror attached to a contact lens which reflected light from a static source onto a moving photographic film[16]. There is no name for this method, but given its similarity to search coils, the author will call it Search Mirrors. Search mirrors were developed solely for the purpose of studying physiological nystagmus, which is movement of the eye when focused on a single static point. This method has not been developed in decades. The most recent and well developed example found in literature review was done by Floyd Ratliff and Lorrin A. Riggs in 1950. Search Mirrors were developed before the existence of digital cameras and personal computers, so a roll of moving photographic film was used instead. A lens was used as shown in figure 10 to make the light from the source hit the mirror in parallel. The beams of light after reflecting off of the mirror would pass through the lens again and converge on the film. For this to work the source of light and the film must be exactly the same distance from the mirror. This requires the user wearing the mirror to be in a fixed position in relation to the light source and film. As can easily be imagined, this basic setup can only record eye movement in one dimension. What is special about Dr. Ratliff and Dr. Lorrin's device is it uses prisms to record eye movements in both dimensions. This system is almost certainly unfeasible for use in the purpose of creating a computer interface. Even if a camera were used instead if a moving film, it would still only be capable of recording a very small range of eye movement. The is fine for studying physiological nystagmus when the eye is almost completely still, but it is not enough for a computer interface when the eye's whole range needs to be recorded.

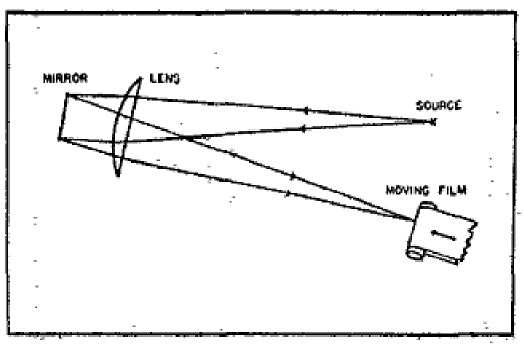

FIG. 1. Basic elements of the recording system

Figure 10. Partial Search Mirror Apparatus [16]

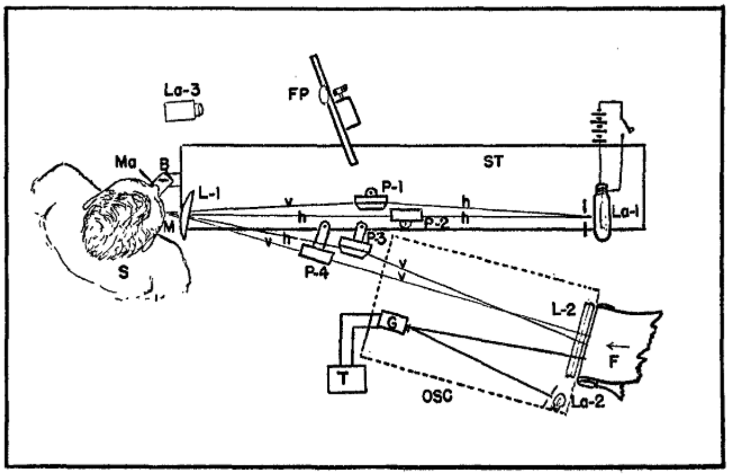

Fig. 2. Diagram of the recording apparatus. B: Biting board. F: Moving film. FP: Fixation point. G: Galvanometer. h and v: h indicates that image of the filament would be horizontal if the rays were brought to a focus at this point, v indicates that the image would be vertical. La-1: Source of light for the recording system. La-2: Oscillograph lamp. La-3: Microscope lamp. L-1: Main lens of recording system. L-2: Cylindrical camera lens. M: Mirror in contact lens. Ma: Mask covering left eye. OSC: Oscillograph. P-1 and P-3: Dove prisms rotated 45° from 'normal' position. P-2 and P-4: Dove prisms in 'normal' position. S: Subject. ST: Steel 'I' beam. T: Timing device.

Figure 11. Whole Search Mirror Apparatus [16]

2.5 Cameras

Naturally, there have been attempts already to track eye gaze using cameras [9,10,11,12,14,15]. The standard camera that comes to mind for most people measures the presence of red, green, and blue and this has been heavy used in development of eye trackers. There is one more type of camera that has been used to great success for the purpose of eye tracking. This camera does not detect light detectable to humans. Instead it tracks infrared light. The different methods using these two cameras are mostly the same. The both use image detection software in the same way to detect eyes. There is generally only one method for using infrared cameras, while there are two types of methods for using standard cameras which are either to have the camera(s) mounted on a headset or have a camera mounted on a computer monitor. In the infrared camera method an infrared light is shined into the user's eyes and the light reflected light gives enough additional information to allow for accurate eye tracking. The idea for this is displayed well in Figure 9. The purpose of using Infrared light for this is that regular light would blind the user while infrared light in invisible to people. The drawbacks to this method are the high cost of the equipment running several thousands of dollars, the accuracy is severely diminished in daylight, and it appears to not be accurate enough to be using in a computer interface. Tobii Glasses is such a system and has been very successful as a market research tool.

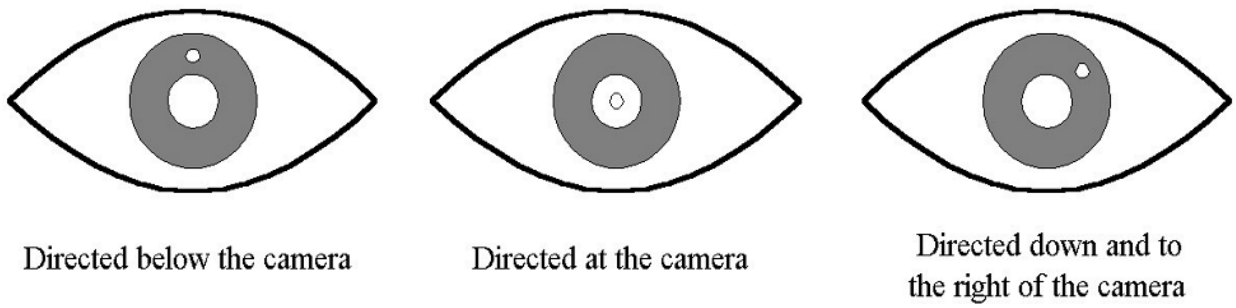

Figure 12. Infrared Camera Image Example [9]

Methods using standard cameras benefit from much cheaper equipment costs, though a completely reliable method has not yet been developed. A popular method is to create a headset with cameras mounted on it and pointed directly into the eyes. The cameras are placed as close to the eyes as possible without obscuring the vision too much. With the head fixed to the head it is possible to know exactly where the eyes are in relation to the headset. However, this does not account for how the eyes or headset are located in relation to the computer screen. This requires the system to be calibrated before each use, forcing the user to remain perfectly still during the session or else the interface will lose accuracy. Dr. Aldo Faisal's lab has created a working example of this headset successfully using it as a computer interface. He expressed an interest in creating a commercial product of this technology, but that was years ago and he has not published any significant progress since then. The other general method is to have the camera mounted near the computer screen. There are many variations of this method, though none appear to be accurate and fast enough to be developed into a viable computer interface.

2.6 Brain Computer Interfaces(BCI)

BCI is an old technology that first appeared in 1977 when The Pentagon's Advanced Research Projects Agency(DARPA) used an electroenphalogram(EEG) to show that brain signals can be used to guide a cursor through a 2D maze [3]. Although this was the first appearance of a true BCI, before this EEG technology and speculation as to its potential as a signal source to control devices had been in existence for decades. EEG is when an array of electrodes is placed either on or under the scalp. Alternatively there are Brain-Machine Interfaces(BMI) which is when an electrode is placed directly on the scalp. Obviously EEG is far less invasive and dangerous, however the signal recovered from BMI is vastly superior to EEG.

2.6.1 Brain-Machine Interface(BMI)

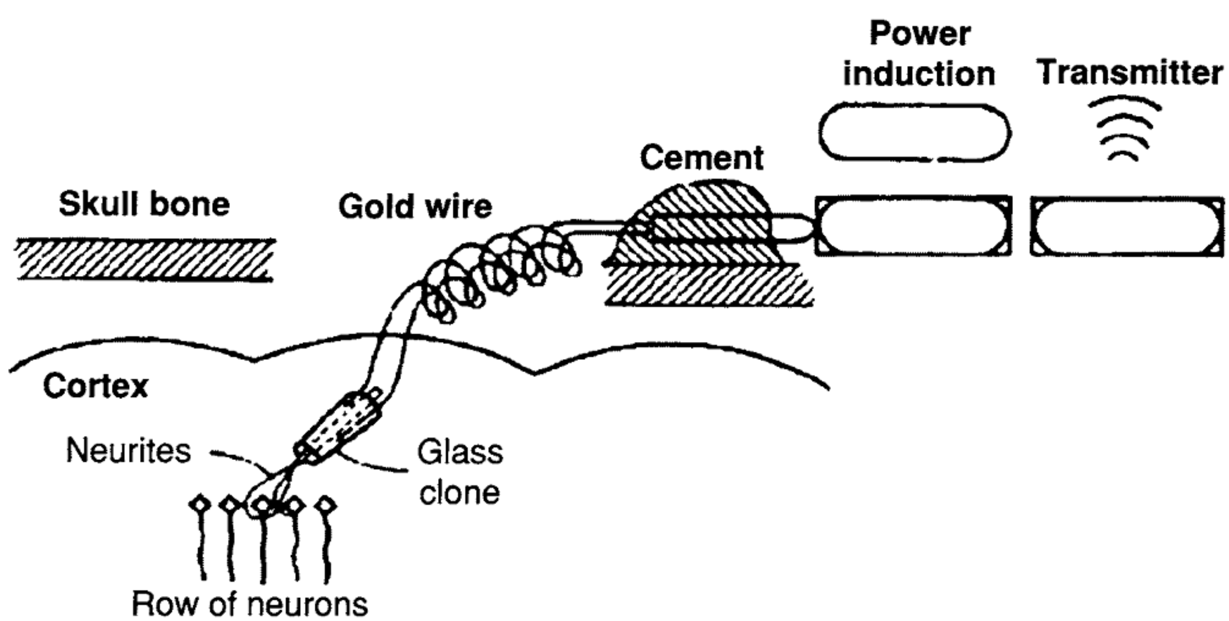

Figure 13. Brain Electrode Implant [3]

The basic idea behind BMI is very simple. A wire is place on a neuron to extract an electrical signal from that neuron. If more neurons are probed then the control of the machine will be greater. That signal is transferred either by wire or through radio. The signal from a single neuron is also very simple. it is either active or not active. If a neuron is active its electrical signal will be high. If the neuron is inactive then its signal will be low. To set a simple machine control of either having the machine off or on there merely needs to be a threshold set so that the machine turns on when the neuron's signal goes over that threshold. For more complicated control more neurons have to be probed. At first the brain will not be tuned to control the machine, which means the user won't have control over the machine. But given the nature of the brain the user only needs to practice to tune the brain and gain control. Not just any neuron will do, in specific the electrode needs to be placed on a neuron responsible for motor control. There is a strong understanding of the location of the motor control neurons, however the understanding of how those neurons direct movement of limbs is very limited. What is understood is that control comes from not a single neuron, but many of them. It should also be noted that the motor control center of the brain rewrites itself after a spinal cord injury, which means that the brain implant will stop working overtime in the people that need these implants most. Given how invasive brain implants are, the need to return and adjust the implant need to be kept to a minimum. To improve the quality of BMI technology it will be necessary to gain a better understanding of the motor control center of the brain.

2.6.2 Electroenphalography(EEG)

Electroenphalography uses an array of electrodes typically placed on the scalp to record signals from the brain. Unlike BMI which receives a clear signal from a single neutron or a small cluster of neurons, EEG records everything. This creates a huge amount of data all jumbled together which has to be separated to find a usable control signal. The first step to separating this data is to understand what different sort of signals there are.

2.6.2.1 Different Types of Signals

Different frequency bands are named with delta band being 0.5-3 Hz, theta band being 4-7 Hz, alpha band being 8-13 Hz, beta band being 14-30 Hz. Most BCI research is in the alpha and beta bands. Several of the signals that can be derived from EEG recordings are oscillatory EEG activity, event-related potentials (ERPs), event-related synchronization/desynchronization (ERS/ERD), visual-evoked potentials, P300, and slow cortical potential.

Oscillatory EEG Activity is when a group of neurons forms a feedback loop, forming an observable frequency. This frequency decreases when the number of neurons in the feedback loop increases. Two such feedback loops are the Rolandic mu-rhythm, in the 10-12 Hz range, and the central beta rhythm, in the 14-18 Hz range, which both appear continuously when the mind is at rest and change in frequency and amplitude when the mind is not at rest.

Event-Related Potentials (ERPs) are signals that occur a fixed time after an external or internal event such as sensory stimulation, specific thoughts, or the absence of an otherwise continual stimulus. An event such as a flash of light or sudden sound is external and such signals are called Exogenous ERP. When someone is spelling words and that person spells with the letter R a signal will be generated moments later different from the signals corresponding to other letters.

An event such as spelling is an internal event and the signals generated are called Endogenous ERP.

Visual-Evoked Potentials(VEP) are EEG signals that appear in response to visual stimulus. It is necessary to be aware that signals generated from movement of eye muscles can be confused for this type of signal. A commonly used type of VEP is the steady-state visual evoked potential(SSVEP).

2.6.2.2 Training

To control a BCI the user must reliably and consistently control their EEG activity. To do this they must understand how their thoughts influence their EEG activity. This is harder to learn than motor functions because unlike moving an arm or a leg, there is no natural form of feedback for the user to learn from. The user can only learn from watching the response of the components of the BCI system. Therefore, the method of training is determined be the signal chosen to control the BCI component. There are two general methods of training: cognitive tasks and operant conditioning. Cognitive tasks is when the EEG signal from a particular mental activity, such as counting continuously, is used as the input to the BCI component. Cognitive tasks have the advantage of being easy to learn but have the disadvantage of not controlling the BCI component directly. Operant conditioning uses a unique signal not previously corresponding to any of the user's cognitive tasks to control the BCI component directly. This takes more time to learn however, and requires the BCI component input to be any thought at all. Gradually, more and more signals are removed as input signals to the BCI component until only the chosen signal remains. Given the dynamic nature of brain signals, it is necessary to observe how different conditions will affect the BCI signal and influence training. Some of these conditions are concentration, distractions, frustration, emotional state, fatigue, motivation, and intentions.

2.7 Conclusion

There appears to be only two types of eye tracking similar to mine in the market place. The first and most successful is the infrared camera method used by Tobii Glasses [10]. This method although quite commercially successful has not been used as part of a computer interface. It is marketed to market researchers who want to study what people look at when they see ads. The second eye tracker has been developed by Dr. Aldo Faisal [11]. Although he has published his findings and his work is for the express purpose of creating a computer interface for the disabled, there is no mention of a computer interface from him being sold to the public. Although there is research, it seems that there is no serious attempt to market a eye tracker based computer interface. Even if Tobii Glasses or a brand like it was modified to work as input for a computer interface, it would still cost thousands of dollars and most likely be too inaccurate to truly be viable. As for Dr. Faisal's system, it is cumbersome requiring the user to have cameras placed directly in front of their face, also requiring the user to lay down and not move. Even if it were to become commercially available it would have significant drawbacks. There is still a space to be filled which is an accurate computer monitor mounted camera eye tracker, which is the aim of this proposal. Such a system if it works would be comparatively very cheap and convenient, circumventing the problems of the current products on the market.

Chapter 3 System Design and Validation

3.1 Introduction

The idea is to use a laptop with a built in camera and a pair of circular glass frames without the glass. Harry Potter Halloween costume glasses work great for this. The user of the HCI would wear the glasses and a program on the computer would convert video from the built in camera into coordinate points of where the user's eyes are looking on the computer screen. These coordinate points would be used to direct where the cursor would be on the computer screen. What is special about this system is that the circular glasses provide additional data to work with which helps to build a frame of reference.

The idea behind the frame of reference is that the location of the eye can be found in relation to the glass frame, the location of the glass frame can be found in relation to the camera, and the location of the camera can be found in relation to the computer screen. With those three references combined it should be possible to find the location of the eye in reference to the computer screen. From there it should only take a little simple trigonometry with the calculated angle of the eye to find where the user is looking on the computer screen. To outline, there are four individual parts of this model: the center of the eye in relation to the center of the glass frames, the center of the glass frames in relation to the center of the camera's vision, the center of the camera's vision in relation to the center of the computer screen, and the angle of the eye's position in relation to the camera.

3.2 Eye Location in Reference to the Glass Frames

At least at the beginning of development it will be necessary for the glass frames to be securely fastened to the user's face so that the frames do not shift. It is possible that the glass frames are already secure with just their basic design. It is uncertain how exactly the location of the center of the eye will be found in relation to the center of the glass frames. Measurements can be taken of the distance from the surface of the eye to glass frames, however the inside of an eye cannot be easily measured. It may be possible to find data of eye anatomy that will have the dimensions recorded. Even then however the measurements will vary for each user and will have to be calibrated for each user. Perhaps the user can be directed to look at particular points of the screen and the resulting data can be used to perfectly calibrate this data.

3.3 Glass Frames Location in Reference to the Camera

For the sake of a simplified beginning it will be assumed that the glass frames are perfectly perpendicular to the camera so that they are viewed as being circular by the camera. There are 2 variable sought to be determined here: how far the center of the lenses is from the camera, and the angle at which the lenses are offset from the center of the camera's vision. The distance is found from two measured pieces of data: the diameter of the physical circular lenses in meters, and the diameter of the virtual circular lenses in pixels. Experiments with these two pieces of data should generate an equation which discovers the distance between the lenses and camera bases on the diameter of the image of the lenses. Such experiments might be to take an image of the lenses when they are 1 meter away and 5 meters away. How to find the angle of the offset of the lenses from the center of the camera vision is unknown. Such an equation would be entirely a property of the camera. That knowledge is surely held in the field of digital camera development. With the distance and the angle it should only take very simple trigonometry to find the exact position of the circular frame in relation to the camera. In the event that the frames are tilted they would appear oval-shaped. This would complicate calculations, but not much. The greatest width of the oval is equivalent to the diameter of its circular form, and the center of the lenses would still be in the same place.

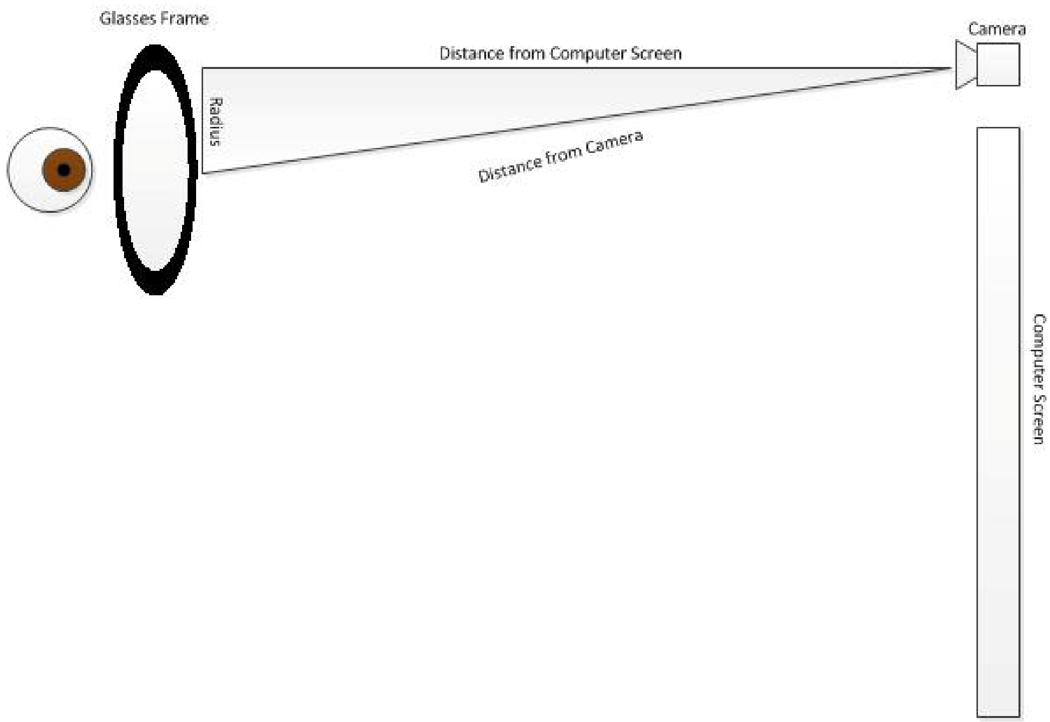

Figure 14. Side View

3.4 Camera Location in Reference to the Computer Screen

The Camera will need to be fixed in relation to the computer screen, such as with a laptop camera. Such a camera's location can be measured once and from then on will always be known with no further calculation.

3.5 Eye Angle

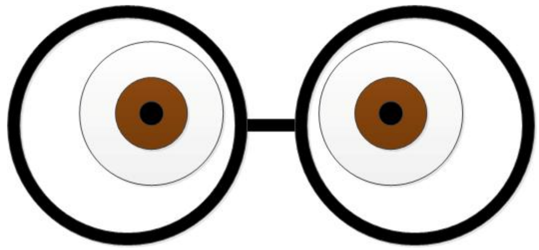

Figure 15. Camera POV

Image recognition software is used to process the image. This will return three pieces of data: The center of the image of the eye, the center of the image of the circular glasses, and the diameter of the image of the circular glasses. The first step to acquiring the angle is to record where the center of the lenses and the center of the eyes are when the user is looking directly at the camera. This is detailed in Figure 16.

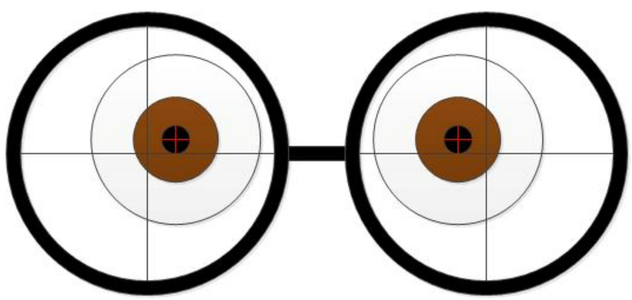

Figure 16. Camera POV with Marked References

The center of the eye on the image when the eyes are looking directly at the camera is the origin. The origin is the place where the angle is zero degrees in the left-right and the up-down directions. When the eyes are looking some place other than the camera, such as in figure 17, then some frame of reference is needed to know where the origin is. Many researchers have tried to use the face as a frame of reference by using facial recognition. However, the accuracy of facial recognition is limited. This is what the purpose of the glasses is. The origin can be found in reference to the center of the circular frames.

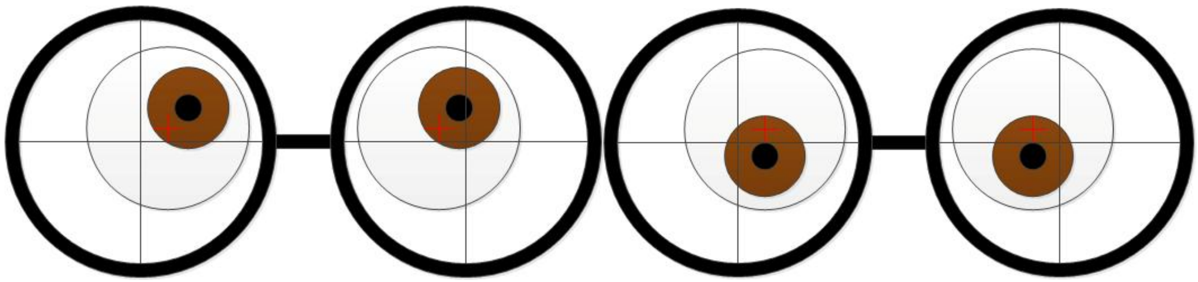

Figure 17. Camera POV Examples

When the eye is not looking directly at the camera as in Figure 17, the location of the center of the image of the pupil in comparison to the location of the origin and the distance between the lenses and the camera should be enough information to determine the angle at which the eye is rotated off of the origin.

3.6 Review

It has been generally shown that position of the eye can be found in relation to the glass frames, the location of the glass frames can be found in relation to the camera, and the camera can be found in relation to the computer screen. This combined means that the eye location can be found in relation to the computer screen. It has also been shown that the angle of the eyes in relation to the camera can be found. The location and orientation of the eyes in relation to the computer screen can be put through some simple trigonometry which will find a pinpoint location on the computer screen as to where the eyes are directed to. This data can be used to directs a computer cursor, which is the primary goal of this project.

3.7 Results

What has been accomplished so far is a program has been created in NI Vision Assistant that finds the center of the glass frames. In figure 18 it is seen that an image is processed and a marker is found. A white marker is found instead of the bare black frames because it was found that black objects are much harder to find that white objects. Figure 19 shows the center of the glass frame on the image. Upon inspection this data was found to be correct.

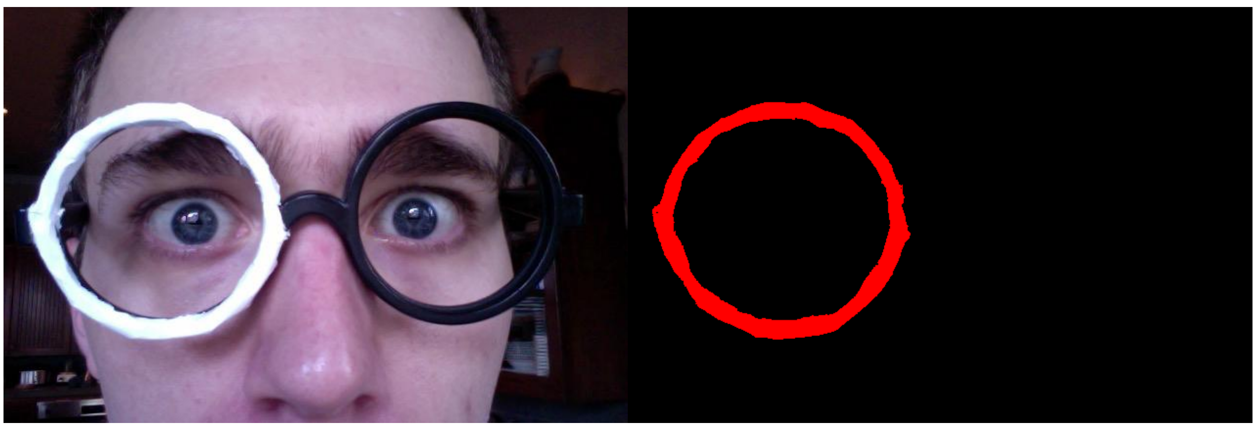

Figure 18. Lens Detection

Results ...	1						
Center of Mass X	159.71258						
Center of Mass Y	222.47651						
First Pixel X	140.00000						
First Pixel Y	98.00000						
Bounding Rect Left	26.00000						

Figure 19. Lens Center Detection

3.8 Discussion

The program made has not been rigorously tested and most likely is not very robust, however it proves that the data needed from the glass frames can be found. The program does not find data about the eyes, which it will need to do before the necessary cursor data can be found. There is open source C++ library called OpenCV which can be used for eye detection. When the project continues it should be done in C++ so it can use the OpenCV library. Also, it should be noted that the ultimate goal is not just to track eye gaze, it is to create a comfortable computer interface. Making an interface that people will want to use, and will be available on all the different operating systems will have its own challenges that has not been discussed here.

Chapter 4 Conclusions and Future Work

It is believed that an image of a person facing the camera wearing glasses can be processed to determine where the person is looking in relation to the camera. This could be a cheap, reliable way for a paralyzed person to interface with a computer by controlling the computer cursor using their eye gaze. The work accomplished so far recovers data from the glass frames. Eye data has not been recovered in this study, but it has in other research so it is proven to at least be possible.

Future work includes writing a program that can find visual data about the eyes. The theory explained in chapter 3 also needs to be implemented or adjusted to process the raw data into eye gaze data. After this software will have to be written that will control the computer cursor at the direction of the eye gaze data. This software will have to include some sort of calibration to improve accuracy for new users.

References

[1] Ayres, T. and Nolan, B., 2006, Voice activated command and control with speech recognition over WiFi, Science of Computer Programming, Vol. 59, No. 1-2, pp. 109-126.

[2] Betke, M., Gips, J., and Fleming, P., March 2002, The Camera Mouse: Visual Tracking of Body Features to Provide Computer Access for People With Severe Disabilities, IEEE Transactions On Neural Systems And Rehabilitation Engineering, Vol. 10, No. 1.

[3] Vallabhaneni, A., Wang, T., and He, B., 2005, Brain-Computer Interface, Springer.

[4] Ebrahimi, T., Vesin, J., and Garcia, G., January 2003, Brain-Computer Interface in Multimedia Communication, IEEE Signal Processing Magazine, pp. 14-24.

[5] Barea, R., Boquete, L., Lopez, E., and Mazo, M., July 1999, Guidance of a Wheelchair Using Electrooculography, Proc. of the 3rd IMACS International Multiconference on Circuits, Systems, Communications and Computers (CSCC'99)

[6] Gips, J. and Olivieri, P., March 1996, Eagle Eyes: An Eye Control System for Persons With Dissabilities, The Eleventh International Conference on Technology and Persons with Disabilities

[7] Robinson, D., October 1963, A Method of Measuring Eye Movement Using a Scleral Search Coil in a Magnetic Field, IEEE Transactions on Bio-Medical Electronics, Vol. 10, No. 4, pp. 137-145.

[8] Kenyon, R., May 1985, A Soft Contact Lens Search Coil For Measuring Eye Movements, Vison Res., Vol. 25, No. 11, pp. 1629-1633.

[9] Poole, A., Ball, L., 2003, Eye tracking in Human-Computer Interaction and Usability Research: Current Status and Future Prospects, Encyclopedia of Human Computer Interaction, Idea Group Reference.

[10] Brône, G., Oben, B., Beeck, K., and Goedemé, T., September 2011, Towards a more effective method for analyzing mobile eye-tracking data: integrating gaze data with object recognition algorithms, UbiComp

[11] Abbot, W. and Faisal, A., 2012, Using ultra-low cost 3D gaze tracking as an intuitive, non-invasive, high information throughput alternative to BMIs, Front. Comput. Neurosci. Conference Abstract: Bernstein Conference 2012.

[12] Kim, E. and Park, S., September 2006, Computer Interface Using Eye Tracking for Handicapped People, Intelligent Data Engineering and Automated Learning – IDEAL 2006, Springer.

[13] Patmore, D. and Knapp, R., 1998, Towards an EOG Based Eye Tracker for Computer Control, In Proceedings of the 3rd international ACM conference on assistive technologies, ASSETS'98.ACM, New York, pp 197–203.

[14] Li, D., Babcock, J., and Parkhurst, D., 2006, OpenEyes: a Low-Cost Head-Mounted Eye-Tracking Solution, Association for Computing Memory. ACM Press, New York, pp. 95–100.

[15] Magee, J., Scott, M., Waber, B., and Betke, M., 2004, EyeKeys: A Real-time Vision Interface Based on Gaze Detection from a Low-grade Video Camera, Computer Vision and Pattern Recognition Workshop

[16] Ratliff, F. and Riggs, L., December 1950, Involuntary Motions of the Eye During Monocular Fixation, Journal of Experimental Psychology, Vol. 40, No. 6, pp. 687-701.

[17] Bulling, A., Ward, J., Gellersen, H., and Troster, G., December 2011, Eye Movement Analysis for Activity Recognition Using Electrooculography, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 33, No. 4, pp. 741-753.